Optimal Power Allocation for Parallel Gaussian Broadcast Channels with Independent and Common Information

Nihar Jindal and Andrea Goldsmith
Department of Electrical Engineering
Stanford University
E-mail : njindal,andrea@systems.stanford.edu

Abstract— **We consider a set of parallel, two-user Gaussian broadcast channels, where the transmitter wishes to send independent information to each of the receivers and common information to both receivers. The capacity region of this channel has been implicitly characterized in the past, but we provide an explicit characterization of the power and rate allocation schemes that achieve the boundary of the three-dimensional rate region. Unlike the broadcast channel with only independent information, we find that the optimal power allocation policy cannot be viewed as a generalization of single-user water-filling. We also consider MIMO broadcast channels, which are non-degraded in general. We propose an achievable region based on dirty paper coding, and discuss the maximum common information rate achievable over these channels.**

I. Introduction

As wireless networks evolve, it is apparent that multi-cast (i.e. sending a common message to all users on a downlink channel) is an important mode of communication that systems will require in the future. In cellular networks, for example, multicast information could be common information such as news updates or location-based information. It is reasonable to assume that networks will want to transmit a mixture of common information to all users and independent information to each of the users. With this in mind, we consider broadcast channels with both common and independent information.

We consider parallel two-user Gaussian broadcast channels, where the transmitter wants to send independent information to users 1 and 2 at rates R_1 and R_2, respectively, and common information (decodable by both users) at rate R_0. For degraded broadcast channels, the common information rate and the independent information rate to the degraded user are interchangeable, because the strongest user can decode anything that the degraded user can. However, we consider parallel channels where in some channels User 1 is the degraded user, but in other channels User 2 is the degraded user. The capacity region of this channel (for both discrete memoryless channels and for Gaussian channels) was characterized in [1] in terms of a union of regions, where the union was taken over different power distributions between the different channels. We first derive an equivalent expression for this capacity region that is more amenable to optimization techniques. We then pose the problem of characterizing the optimal power and rate allocation schemes that achieve the boundary of the three-dimensional region using Lagrangian techniques. We then apply the utility function approach used for the broadcast channel [2] without common information, but we find that this approach does not work in general. We use a more direct approach to maximize the Lagrangian function and obtain the capacity region with common information using this approach. Using this method, the optimal allocation is found by performing a finite maximization in each channel.

Finally, we consider MIMO broadcast channels, which in general are not degraded. Thus, the capacity region with or without common information is not known for this channel. We propose an achievable region based on dirty paper coding. We also consider the maximum common rate achievable on these channels, i.e. the common information capacity.

The remainder of this paper is organized as follows. In Section II we describe the system model, followed by the capacity region of the broadcast channel in III. In Section IV we describe the Lagrangian formulation used to find the optimal power allocation, along with a method to maximize the Lagrangian. In Section VI we describe a simple procedure to find the optimal Lagrangian multipliers, followed by some numerical results in VII. We briefly discuss MIMO broadcast channels in VIII, followed by our conclusions.

II. System Model

We consider the following channel:

$$y_1(i) = x(i) + z_1(i) \quad i = 1, \ldots, N \tag{1}$$
$$y_2(i) = x(i) + z_2(i) \quad i = 1, \ldots, N \tag{2}$$

where $z_1(i) \sim N(0, N_1(i))$ and $z_2(i) \sim N(0, N_2(i))$. If $N_1(i) \leq N_2(i)$ for all i, then this is a *degraded* broadcast channel. For such a channel, common information and independent information sent to User 2 are interchangeable, and the optimal power allocation is essentially equivalent to that for channels with only independent information [2, 3]. We will only consider the non-degraded case, i.e. where for some i we have $N_1(i) \leq N_2(i)$ and for some other i we have $N_2(i) \leq N_1(i)$. We impose an average power constraint P on the input, i.e. $\sum_{i=1}^{N} E[x(i)^2] \leq P$.

III. Capacity Region Characterization

In [1], the capacity region for two parallel Gaussian broadcast channels (non-degraded) with a common message and independent messages for both users is given. This characterization can easily be extended to N parallel channels

Theorem 1: The capacity region of N parallel two-user broadcast channels is equal to the convex hull of the union of the regions defined by

$$\begin{aligned}
R_0 &\leq \min\left(\sum_{i=1}^{N} C\left(\frac{P_0(i)}{P_1(i)+P_2(i)+N_1(i)}\right),\right. \\
&\qquad \left.\sum_{i=1}^{N} C\left(\frac{P_0(i)}{P_1(i)+P_2(i)+N_2(i)}\right)\right) \\
R_1 &\leq \sum_{i\in A_1} C\left(\frac{P_1(i)}{N_1(i)}\right) + \sum_{i\in A_2} C\left(\frac{P_1(i)}{P_2(i)+N_1(i)}\right) \\
R_2 &\leq \sum_{i\in A_1} C\left(\frac{P_2(i)}{P_1(i)+N_2(i)}\right) + \sum_{i\in A_2} C\left(\frac{P_2(i)}{N_2(i)}\right)
\end{aligned}$$

where the union is taken over all $P_0(i), P_1(i), P_2(i)$ such that $\sum_{i=1}^{N} P_0(i)+P_1(i)+P_2(i) \leq P$, and where A_1 is the set of $i \in 1,\ldots,N$ such that $N_1(i) \leq N_2(i)$ and A_2 is the complementary set, i.e. the set of i such that $N_2(i) < N_1(i)$.

Proof: The converse for this region is a straightforward generalization of [1]. Achievability follows from standard arguments similar to [3]. ■

The powers $(P_0(i), P_1(i), P_2(i))$ can be interpreted as the power allocated to send the common message, the independent message to user 1, and the independent message to user 2, respectively. The common message is decoded first (with the powers $P_1(i)+P_2(i)$ treated as interference), followed by the independent messages. For $i \in A_1$, User 1 can decode and subtract out the codeword intended for User 2 before decoding his own codeword. For $i \in A_2$, User 1 must treat $P_2(i)$ as interference.

For transmission of the independent messages, separate codebooks (and rates) are used for each user on each of the N channels. These codewords are decoded *independently* on each channel. However, the common message codebook cannot be broken into different codebooks for each channel, i.e. joint encoding and joint decoding must be performed across the channels to achieve capacity. If the common message was broken into different codebooks for each channel, the common rate transmitted on each channel would be limited by the weakest user in *each* channel (since the stronger user can decode anything that the weakest user can, by degradedness). The corresponding common rate would be given by $\sum_{i=1}^{N} C\left(\frac{P_0(i)}{P_1(i)+P_2(i)+\max(N_1(i),N_2(i))}\right)$ (without any minimization operation required). This is highly sub-optimal, and much higher common information rates can be achieved by jointly decoding. Each user extracts a different amount of information about the common message from each of the channels due to the different noise powers of the users on each channel, and we consider the total amount of mutual information across all channels. This is similar from an information theoretic point of view to a flat-fading single-user channel where only the receiver knows the channel state information. In such a scenario, encoding must be done across all different channel fades and joint decoding must be performed.

IV. Formulation of Optimization

From the previous section, we see that the capacity region can be defined as the convex hull of the union of all rate points described in the previous section, where the union is taken over all power allocations satisfying the average power constraint. Since the capacity region is convex, we can fully characterize it by maximizing the weighted sum of rates, for different weights. We wish to find the optimal power allocation policy that maximizes the weighted sum of rates for arbitrary rates. This is given by the following problem;

$$\max_{(R_0,R_1,R_2)\in C_{BC}(\overline{P})} \mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0 \tag{3}$$

Using standard convex optimization techniques, for the optimal λ, this is equivalent to:

$$\max_{\mathbf{P}(i)} \mu_1 R_1(\mathbf{P}(i)) + \mu_2 R_2(\mathbf{P}(i)) + \tag{4}$$
$$\mu_0(\min(R_{01}(\mathbf{P}(i)), R_{02}(\mathbf{P}(i)))) - \lambda\left(\sum_{i=1}^{N} P(i) - \overline{P}\right)$$

where $R_1(\mathbf{P}(i))$ and $R_2(\mathbf{P}(i))$ are defined as:

$$R_j(\mathbf{P}(i)) = \sum_{i=1}^{N} \log\left(1 + \frac{P_j(i)}{N_j(i) + P_l(i)\mathbf{1}[i \notin A_i]}\right)$$

for $j,l = 1,2$ and $l \neq j$, R_{0j} is defined as:

$$R_{0j}(\mathbf{P}(i)) = \sum_{i=1}^{N} \log\left(1 + \frac{P_0(i)}{N_j(i)+P_1(i)+P_2(i)}\right)$$

for $j = 1,2$, and $P(i)$ is defined as $P(i) = P_0(i) + P_1(i) + P_2(i)$. This maximization can further be simplified by replacing the minimum operation with a weighted sum of the two common rates. It can be shown (using standard convex optimization methods) that the optimal power allocation policy solves

$$\max_{\mathbf{P}(i)} \mu_1 R_1(\mathbf{P}(i)) + \mu_2 R_2(\mathbf{P}(i)) + \lambda_1 R_{01}(\mathbf{P}(i)) + \tag{5}$$
$$\lambda_2 R_{02}(\mathbf{P}(i)) - \lambda\left(\sum_{i=1}^{N} P(i) - \overline{P}\right)$$

for the optimal Lagrangian multipliers $(\lambda, \lambda_1, \lambda_2)$, which each must be non-negative and satisfy $\lambda_1 + \lambda_2 = \mu_0 = (1-\mu_1-\mu_2)$. Furthermore, for any λ, the solution to (4) is equal to the solution to (5) for λ_1 and λ_2 such that the optimizing power allocation yields either $R_{01} = R_{02}$ or $\lambda_i = 0$ for one of the users.

In the next section we describe how to solve (5) for any $(\lambda, \lambda_1, \lambda_2)$. In section VI we describe a simple method to find the optimal Lagrange multipliers.

V. Maximization of Lagrangian

In this section we describe a method to solve (5), i.e. maximize the weighted sum of rates given the power price λ and the Lagrangian's λ_1 and λ_2. First note that a power allocation solves (5) if and only if it is the solution to

$$\max_{P_0(i),P_1(i),P_2(i)} \mu_1 R_1(P(i)) + \mu_2 R_2(P(i)) + \lambda_1 R_{01}(P(i)) \tag{6}$$
$$+\lambda_2 R_{02}(P(i)) - \lambda\left(P_0(i) + P_1(i) + P_2(i)\right)$$

for each $i = 1,\ldots,N$. When there is no common information (i.e. $\lambda_1 = \lambda_2 = 0$), (6) can be solved using an intuitive utility function approach [2, 3]. In Section V-A we show that this approach does not work in general when there is common information, and we instead must use a less intuitive method.

A. Utility Function Approach

In this section we attempt to use utility functions to determine the optimal power allocation. We use the procedure developed in [2], where it was used to find the optimal power allocation without common information. Without loss of generality, we consider states where $N_1(i) < N_2(i)$. We define the following utility functions:

$$\begin{aligned} u_1(z) &= \frac{\mu_1}{N_1(i)+z} - \lambda \\ u_2(z) &= \frac{\mu_2}{N_2(i)+z} - \lambda \\ u_0(z) &= \frac{\lambda_1}{N_1(i)+z} + \frac{\lambda_2}{N_2(i)+z} - \lambda \end{aligned}$$

If we let J^* denote the solution to (6) and J^* is achieved by $(P_0(i), P_1(i), P_2(i))$, then we have

$$\begin{aligned} J^* &= \int_{z=0}^{P_1(i)} u_1(z)dz + \int_{z=P_1(i)}^{P_1(i)+P_2(i)} u_2(z)dz + \\ &\quad \int_{z=P_1(i)+P_2(i)}^{P_0(i)+P_1(i)+P_2(i)} u_0(z)dz \\ &\leq \int_{z=0}^{\infty} \left[\max_i u_i(z)\right]^+ dz \end{aligned}$$

where the argument of the final integral is a pointwise maximum of the utility functions. This upper bound is achievable if the maximum of the utility functions is in order of decreasing channel gains (i.e. $u_1(z)$ is the maximum function initially, followed by $u_2(z)$ and then $u_0(z)$). When there is no common information, then this condition is satisfied and the upper bound is achievable [2]. Thus, the optimum power allocation can be found by taking the pointwise maximum of the utility functions corresponding to both users. However, for common information, the ordering of the utility functions does not always satisfy this condition and thus the utility function approach can not be used in general. There are a number of interesting scenarios where the utility function approach does work:

1. $\mu_0 < \mu_1$ and $\mu_0 < \mu_2$: If this condition is satisfied, then $u_0(z) \leq u_j(z)$ for all $z \geq 0$, where User j is the user with the larger noise power. Thus no common information is transmitted in each state, and this simplifies to the standard independent information BC, for which the utility function approach works.
2. $\mu_0 \geq \mu_1$ and $\mu_0 \geq \mu_2$: In this scenario, $u_0(z) \leq u_j(z)$ for all $z \geq 0$, where User j is the user with the larger noise power. Thus we are left with the common rate user and the better of two users. In this scenario it can be shown that the utility functions are ordered correctly such that the upper bound is achievable. This includes the interesting case when $\mu_0 \geq \mu_1 = \mu_2$.

Interestingly, the utility function approach does not work when either $\mu_1 = 0$ or $\mu_2 = 0$. In Fig. V-A, utility functions are shown for a channel where $N_1 = 1.5$, $N_2 = 1$, $\mu_1 = .54$, $\lambda_1 = .1$, and $\lambda_2 = .36$. For $0 \leq z \leq 1.25$, $u_0(z)$ is the largest utility function, but for $z > 1.25$, $u_1(z)$ is the largest function. Therefore, the upper bound J^* can not be achieved, because User 0 must be allocated power after (i.e. for larger z) User 1 is allocated power. It is also possible to find counter-examples for situations where all three priorities are non-zero.

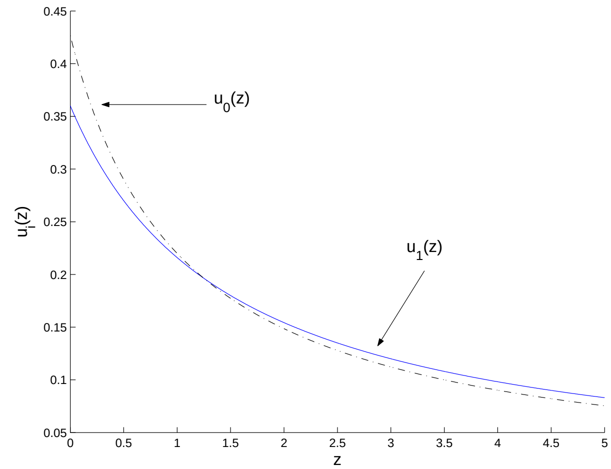

Fig. 1. Utility Functions for a sample channel with $\mu_2 = 0$

B. Direct Approach

Since the utility function approach does not work in general, we must consider a more direct approach to maximize the Lagrangian function. The most straightforward way of maximizing a continuous function is to consider the points where the derivatives of the function are zero. The only complication is that the powers $P_0(i), P_1(i)$, and $P_2(i)$ must all be non-negative. Thus, for $j = 1, 2, 3$, the derivative of the objective function with respect to $P_j(i)$ must be equal to 0 if the optimal $P_j(i)$ is strictly positive. Additionally, if $P_j^*(i) = 0$ for some j, then the derivative with respect to $P_j(i)$ at $P_j(i) = 0$ must be less than or equal to zero. The simple structure of the objective function implies that each partial derivative is equal to zero at only one point. Thus, it is sufficient to consider the 8 different combinations of power allocations. The maximum of the objective function is then equal to the maximum of these 8 possible combinations.

Without loss of generality, consider a state where $N_1(i) \leq N_2(i)$. The partial derivatives of (6) are given by:

$$\begin{aligned} \frac{\partial J}{\partial P_1} &= \frac{\mu_1}{P_1+N_1} + \frac{\mu_2}{P_1+P_2+N_2} - \frac{\mu_2}{P_1+N_2} + \\ &\quad \frac{\lambda_1}{P_0+P_1+P_2+N_1} - \frac{\lambda_1}{P_1+P_2+N_1} + \\ &\quad \frac{\lambda_2}{P_0+P_1+P_2+N_2} - \frac{\lambda_2}{P_1+P_2+N_2} - \lambda \\ \frac{\partial J}{\partial P_2} &= \frac{\mu_2}{P_1+P_2+N_2} + \frac{\lambda_1}{P_0+P_1+P_2+N_1} - \\ &\quad \frac{\lambda_1}{P_1+P_2+N_1} + \frac{\lambda_2}{P_0+P_1+P_2+N_2} - \\ &\quad \frac{\lambda_2}{P_1+P_2+N_2} - \lambda \\ \frac{\partial J}{\partial P_2} &= \frac{\lambda_1}{P_0+P_1+P_2+N_1} + \frac{\lambda_2}{P_0+P_1+P_2+N_2} - \lambda. \end{aligned}$$

Since there are three different powers to be allocated, there are three different sets of partial derivatives to consider. First consider the four cases corresponding to $P_0 = 0$. By setting some of these derivatives to zero (i.e. the partials corresponding to users with non-zero power), we find the optimal allocations are

given by the following when $P_0 = 0$:
1. $P_1 = P_2 = 0$
2. $P_1 > 0,\ P_2 = 0$: $P_1 = \frac{\mu_1}{\lambda} - N_1$
3. $P_2 > 0,\ P_1 = 0$: $P_2 = \frac{\mu_2}{\lambda} - N_2$
4. $P_1 > 0,\ P_2 > 0$: $P_1 = \frac{\mu_1 N_2 - \mu_2 N_1}{\mu_2 - \mu_1},\ P_2 = \frac{\mu_2}{\lambda} - N_2 - P_1$.

When $P_0 > 0$ the power allocations are given by:
1. $P_0 > 0, P_1 = P_2 = 0$: $P_0 = P_{thresh}$
2. $P_0, P_1 > 0,\ P_2 = 0$: $P_1 = \frac{\lambda_2 N_1 - (\mu_1 - \lambda_1) N_2}{\mu_1 - \mu_0},\ P_0 = P_{thresh} - P_1$
3. $P_0, P_2 > 0,\ P_1 = 0$: $P_2 = \frac{\lambda_1 N_2 - (\mu_2 - \lambda_2) N_1}{\mu_2 - \mu_0},\ P_0 = P_{thresh} - P_2$
4. $P_0, P_1, P_2 > 0$: $P_1 = \frac{\mu_1 N_2 - \mu_2 N_1}{\mu_2 - \mu_1},\ P_2 = \frac{\lambda_1 N_2 - (\mu_2 - \lambda_2) N_1}{\mu_2 - \mu_0} - P_1,\ P_0 = P_{thresh} - P_1 - P_2$

where

$$P_{thresh} = \frac{1}{2\lambda}(\mu_0 - \lambda(N_1 + N_2) + \sqrt{(\lambda(N_1 + N_2) - \mu_0)^2 + 4\lambda(\lambda_1 N_2 + \lambda_2 N_1 - \lambda N_1 N_2)}).$$

One of these 8 power allocations is guaranteed to achieve the maximum of (6). Thus we can find the maximum of the Lagrangian in each state by evaluating all eight power allocations, checking for non-negativity of powers, and then choosing the allocation that maximizes the objective.

When there is only independent information, it can be shown that only one of the four cases is feasible for different values of λ (i.e. the space $\lambda > 0$ can be decomposed into four mutually exclusive intervals corresponding to the four different allocations). Thus, a closed form solution for the optimal power allocation can be given in terms of λ. However, no such simplification can be done for the situation when there is common information. Thus, in general, the maximum amongst the eight allocations must be performed.

VI. Optimal Lagrange Multipliers

By the KKT conditions, the solution to the original Lagrangian characterization in (4) for the optimal $(\lambda, \lambda_1, \lambda_2)$ will satisfy the power constraint with equality. It is easy to see that the power allocation solving (4) is a decreasing function of λ. Thus, the optimal λ can be found by solving (4) for different values of λ determined by the bisection method (over λ).

To maximize the Lagrangian function in(4), we work with the simplified maximization in (5), where the minimum is replaced with a weighted sum of the common information rates. It can be shown that for any λ, the solution to (4) is equal to the solution of (5) for λ_1 and λ_2 such that the optimizing power allocation yields $R_{01} = R_{02}$ or $\lambda_i = 0$ for one of the two users. This follows intuitively because for any power allocation that yields $R_{01} \neq R_{02}$, we can reallocate $P_0(i)$ over different channels (without increasing the sum of power) to increase the smaller of the two common rates slightly, and thus increase the objective function. However, this is not possible if the allocation of power $P_0(i)$ is already single-user optimal for the user with the smaller common rate (i.e. no reallocation of $P_0(i)$ increases the common rate of the user with the smaller common rate). This corresponds to the scenario where either $\lambda_1 = 0$ or $\lambda_2 = 0$[1].

[1] In general, the minimum of two concave functions occurs at a point where the two functions meet, unless the minimum of the two functions is equal to the maximum of one of the functions.

Furthermore, it can also be shown that the optimizing R_{01} in (4) is a increasing function of λ_1 and the optimizing R_{02} is a decreasing function of λ_2.

Thus, the following procedure can be used to find the optimal Lagrange multipliers $(\lambda, \lambda_1, \lambda_2)$. First choose an initial positive value for λ. Then repeat the following algorithm:

1. Solve (4) by the following procedure:
 (a) Solve (5) with $\lambda_1 = 0$. If $R_{01} \geq R_{02}$ for the optimizing solution, proceed to Step 2.
 (b) Solve (5) with $\lambda_1 = \mu_0$. If $R_{02} \geq R_{01}$ for the optimizing solution, proceed to Step 2.
 (c) Use the bisection method to find λ_1 such that the optimizing solution of (5) satisfies $R_{01} = R_{02}$.
2. If the solution of (4) exactly meets the power constraint, then exit. Otherwise, if the solution of (4) is strictly larger/smaller than the power constraint, then increase/decrease λ and return to Step 1.

Note that the update of λ can be performed using a one-dimensional bisection method, and Step 1(c) can be performed with a one-dimensional bisection method on λ_1. This procedure is implemented in order to find numerical results in section VII

VII. Numerical Results

In this section we present numerical results on the capacity region of a two-user broadcast channel. In Fig. 2 the capacity region of a two user channel is shown for $N = 2$. In state 1, user 1 has an average SNR of 10 dB and user 2 has an SNR of -10 dB. In state 2, the SNR's of the users are reversed from state 1. Notice that due to the large difference in SNR of the two users, the capacity region when $R_0 = 0$ (i.e. no common information) is far from the straight line segment connecting the maximum single-user rate to users 1 and 2. However, if $R_2 = 0$ (or $R_1 = 0$ by symmetry), the region is quite close to time-sharing between transmitting only common information and transmitting only independent information to User 1. In Fig. 3 the capacity region of a two-user channel is shown where in state 1, user 1 has an SNR of 0 dB while user 2 has an SNR of -10 dB. In state 2, the roles of the users are reversed. We again see that the capacity region is relatively flat in the direction of common information (R_0), which implies that time-sharing between sending common information and independent information comes quite close to the actual capacity-achieving strategy.

VIII. MIMO Channels

In this section we consider multiple-input, multiple-output (MIMO) broadcast channels. Since MIMO broadcast channels are not in general degraded, the capacity region with common and independent information is unknown even for a single constant channel (i.e. $N = 1$). In the following sections we discuss an achievable region, followed by discussion of transmitting only common information over a multiple-input, single-output channel.

A. Capacity Region

An achievable region for the common and independent information MIMO broadcast channel can be established using dirty paper coding [4]. Dirty paper coding was shown to achieve the

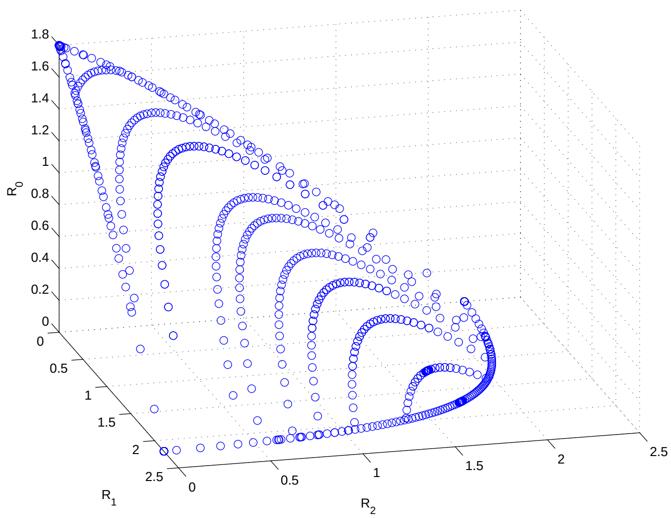

Fig. 2. Plot of capacity region channel with 20 dB SNR difference between the two users

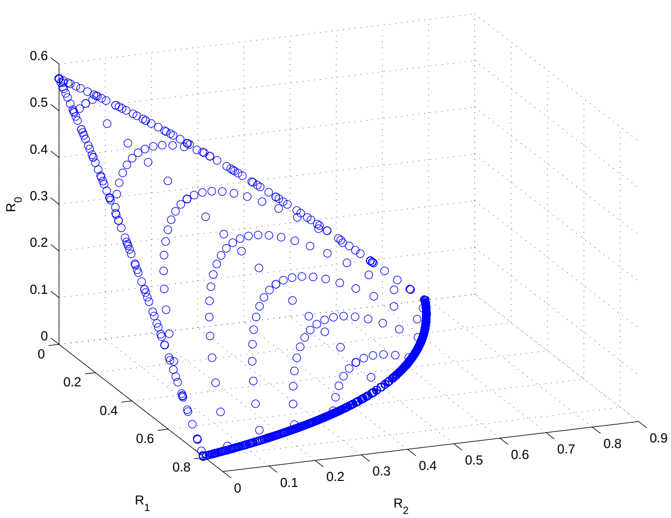

Fig. 3. Plot of capacity region channel with 10 dB SNR difference between the two users

sum rate capacity of the MIMO broadcast channel (i.e. the maximum of $R_0+R_1+R_2$ in the capacity region) in [5–8]. Amraoui et. al. [9] recently considered the rates achievable using successive decoding, a technique that is practically easier to implement than dirty paper coding.

By first encoding the common message followed by the independent messages, the following rate triplet is achievable:

$$R_0 = \min_{j=1,2} \log \frac{\left|\mathbf{I} + \mathbf{H_j}(\Sigma_0 + \Sigma_1 + \Sigma_2)\mathbf{H_j}^T\right|}{\left|\mathbf{I} + \mathbf{H_j}(\Sigma_1 + \Sigma_2)\mathbf{H_j}^T\right|} \quad (7)$$

$$R_1 = \log \frac{\left|\mathbf{I} + \mathbf{H_1}(\Sigma_1 + \Sigma_2)\mathbf{H_1}^T\right|}{\left|\mathbf{I} + \mathbf{H_1}\Sigma_2\mathbf{H_1}^T\right|} \quad (8)$$

$$R_2 = \log \left|\mathbf{I} + \mathbf{H_2}\Sigma_2\mathbf{H_2}^T\right| \quad (9)$$

for any set of positive semi-definite covariances satisfying $Tr(\Sigma_0 + \Sigma_1 + \Sigma_2) \leq P$. Additionally, the ordering of users 1 and 2 can be switched so that user 1 sees no interference and user 2 views Σ_1 as interference. This region can easily be extended to parallel broadcast channels. However, the rate equations given above are not concave functions of the input covariances, and thus finding the boundary region of this achievable region (even for $N = 1$) seems to be difficult from a numerical perspective.

B. Common Information Capacity

Though the above achievable region appears to be difficult to compute, it is far easier to calculate the common information capacity (i.e. the maximum common rate) of a MIMO broadcast channel. Since Gaussian inputs are optimal for MIMO channels, the common rate capacity of a K-use broadcast channel is given by:

$$\mathcal{C}_0 = \max_{\mathbf{\Sigma} \geq 0, Tr(\mathbf{\Sigma}) \leq P} \; \min_{i=1,\ldots,K} \log \left|\mathbf{I} + \mathbf{H}_i \mathbf{\Sigma} \mathbf{H}_i^T\right|. \quad (10)$$

The objective function of this maximization is a minimum of concave functions, and thus is a concave function. Therefore, standard convex optimization techniques can be applied to perform the maximization.

For the case of multiple-input, single-output channels (i.e. single antennas at each of the receivers), it can be shown that a rank-one covariance matrix (i.e. beamforming) achieves the common information capacity when there are two users. Interestingly, beamforming does not in general achieve the common information capacity for more than two users. Consider a system of K unit norm users, each equally spaced around the unit circle. For any $\epsilon > 0$, we can find large enough K such that for any choice of a direction vector $\mathbf{v}$, $\min_{i=1,\ldots,K} |\mathbf{H}_i\mathbf{v}| < \epsilon$, since any direction is nearly orthogonal to at least one user because users are equally spaced around the unit circle. The common rate when using covariance $\Sigma = \frac{P}{M}\mathbf{v}\mathbf{v}^T$ is equal to $\min_{i=1,\ldots,K} \log(1 + \frac{P}{M}\mathbf{H}_i\mathbf{v}\mathbf{v}^T\mathbf{H}_i^T)$. Thus, using beamforming, the common rate goes to 0 as $K \to \infty$. However, by using an identity covariance, i.e. $\Sigma = \frac{1}{M}\mathbf{I}_M$, the mutual information of *each* user is $\log(1+\frac{1}{M}||\mathbf{H}_i||^2) = \log(1+\frac{1}{M})$ which is independent of K. Thus, common information capacity is not achieved by beamforming for large enough K. We have seen this to be true in general for $K > 2$, but there are exceptions for which beamforming does achieve capacity.

IX. Conclusion

Broadcast channels have been heavily studied by information theorists during the past three decades. However, the overwhelming majority of work has concentrated on only independent information. In this paper we considered Gaussian broadcast channels with both independent and common information rate. We first recast the expression for the capacity region in a more traditional manner, and found the optimal power and rate allocation policies that achieve the boundary of the capacity region. Interestingly, the simple approaches that worked in the absence of common information no longer work in general when common information is added to the picture. However, some intuition can still be gleamed from the optimal power allocation policy. Finally, we considered MIMO broadcast channels and proposed an achievable rate region based on dirty paper coding.

REFERENCES

[1] A. E. Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Information Transmission*, vol. 6, no. 1, pp. 3–23, Jan-March 1980.

[2] D. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," unpublished. Available at http://www.eecs.berkeley.edu/ dtse/pub.html.

[3] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels–Part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 1083–1102, March 2001.

[4] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[5] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, July 1999.

[6] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[7] W. Yu and J. M. Cioffi, "Sum capacity of a Gaussian vector broadcast channel," in *Proceedings of Int. Symp. Inform. Theory*, June 2002, p. 498.

[8] P. Viswanath and D. N. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1912–1921, June 2003.

[9] A. Amraoui, G. Kramer, and S. Shamai, "Coding for the mimo broadcast channel," in *Proceedings of IEEE Int. Symp. Inform. Theory*, June 2003, p. 296.